Cloud-scale SDN Network Security in TeraFlow

Alberto Mozo*, Stanislav Vakaruk*, Antonio Pastor†, Rahul Bobba‡, Carlos Natalino§, Marija Furdek§, Raul Muñoz¶, Ramon Casellas¶, Ricardo Martínez¶, Juan-Pedro Fernández-Palacios†, Ricard Vilalta¶

*Universidad Politécnica de Madrid, Madrid, Spain
†Telefónica I+D, Madrid, Spain
‡NEC Labs Europe, Heidelberg, Germany
§Chalmers University of Technology, Gothenburg, Sweden
¶Centre Tecnològic de Telecomunicacions de Catalunya (CTTC/CERCA), Castelldefels, Spain

***Abstract*—TeraFlow proposes a new type of secure, cloud-native Software Defined Networking controller that will radically advance the state-of-the-art in beyond 5G networks by introducing novel micro-services architecture and providing revolutionary features for flow management and optical/microwave network equipment integration. Two key contributions of this project will be the incorporation by design in the TeraFlow controller of (i) security using advanced Machine Learning (ML) techniques both in the data and control planes and (ii) forensic evidence for multi-tenancy based on Distributed Ledgers. This paper presents the TeraFlow Cybersecurity scenario, which will be used to evaluate some of the proposed contributions.**

I. Introduction

Nowadays, when an operator moves towards an automated environment, security becomes key as network operations have to be done by software components virtually operating without human intervention or oversight. In particular, security is one of the most critical aspects that a state-of-the-art Software-Defined Networking (SDN) controller needs to provide [1]. The aim of TeraFlow [2] is to create a novel SDN controller for beyond 5G networks that integrates current Network Function Virtualization (NFV) and Mobile Edge Computing (MEC) frameworks and supports novel features for flow management and equipment integration. Regarding that the evolution of security threats, many of them arising from technological advancements that create new attack vectors such as NFV [3], requires protection of both network services and the network controller from attacks, TeraFlow will address these needs from several perspectives, incorporating distributed and centralized solutions.

To detect malicious flows at the data plane, TeraFlow will devise a distributed Machine Learning (ML)-based Intrusion Detection Systems (IDSs) deployed at the network edge. Many IDSs have been proposed for identifying different types of attacks [4], but there is a lack of proposals addressing scalability and latency problems. A distributed IDS is expected to improve scalability and response time to detect malicious flows, while reporting to the centralized controller for a holistic network security assessment. The centralized cybersecurity component developed by TeraFlow will perform a security assessment of end-to-end physical channels (e.g., lightpaths in the optical layer) to detect malicious attacks that can target the network infrastructure [5]. The service-layer security assessment reports received from the edge IDSs merged with information from the infrastructure layer will enable cross-layer security assessment.

Deploying ML models at the network edge in a distributed fashion implies that less resources are required per edge node, while maintaining the same accuracy and performance level. Recently, there has been an explosion of research interest in Automatic Machine Learning, with special attention on Neural Architecture Search (NAS) [6] that aims to generate a compact but robust and well-performing neural architecture by selecting and combining different basic components from a predefined search space. In this context, TeraFlow proposes "green AI approaches" and the deployment of ML-based threat detectors in resource-limited nodes (e.g. when deployed in P4 switches) by utilizing AutoML techniques to reduce model complexity, while maintaining threat detector performance and effectiveness with respect to the original models. Furthermore, and due to the lack of publicly available attack datasets for training and testing purposes, TeraFlow will apply Generative Adversarial Networks [7] for synthetic traffic generation in order to obtain generators that are specifically tailored to produce synthetic network attacks [8].

Recent studies show that popular ML algorithms, and in particular deep neural networks, have been found to be vulnerable to malicious and well-designed attacks than can easily fool a Deep Learning model with small perturbations imperceptible to humans [9]. Therefore, the so-called adversarial attacks could mislead the ML-based components running on SDN controllers and cause harmful situations in security-critical areas of the network. Although these sophisticated adversarial attacks are still premature, the main conclusion is that testing in training processes is insufficient because it provides a lower bound on the failure rate of the system, and therefore in order to provide security guarantees, an upper bound is necessary. Teraflow will explore the application of novel techniques to provide resiliency to SDN ML-based components against adversarial attacks.

The key features of blockchains, namely, decentralization, immutability, and transparency make the use of blockchains also appealing for managing resources and services in multi-tenant networks. In a nutshell, a blockchain serves here as a database or log that stores, e.g., the allocations of network resources by the various network tenants. In particular, the use of blockchains replaces centralized network management with conventional database management systems. Major advantages are the elimination of trusted third-parties that maintain the databases with single point of failures, and data provenance including data immutability and traceability. Both are cornerstones for a resilient and trustworthy platform for storing and processing sensitive data. Furthermore, smart contracts provide a universal basis to automate, simplify, and secure network management tasks that involve possibly sensitive data from multiple stakeholders of the network. Trust and multi-tenancy are improved in the SDN controllers by introducing novel security mechanisms through the usage of smart contracts and secure consensus algorithms.

In the following section, we detail each one of the components that will be developed within TeraFlow for the security assessment of networks.

II. TeraFlow Network Security

A. Cybersecurity

To enable an encompassing security assessment, TeraFlow will implement 3 use cases covering a varied set of threats. Figure 1 gives a high-level overview of the TeraFlow architecture and the attack vectors that can be exploited by entities to perform malicious activities. The use cases have been grouped around the network planes visible to TeraFlow.

At the data plane, the TeraFlow OS will deal with both classical and advanced network attacks (e.g., DDoS, malware, etc.). This use case will demonstrate that novel approaches enabled by machine and deep learning techniques will allow TeraFlow OS face new threats such as detection of malicious encrypted traffic (e.g., cryptominig malware). As detection and identification of malware network flows crossing the data plane cannot be done in a ML-based central component due to scalability problems and slow response times, we propose to implement a distributed solution in which ML components are deployed at 5G edge nodes. To this end, a feature extractor is deployed at the edge of the network to collect and summarise packets. The flow statistics aggregated by the feature extractor are sent to a ML classifier. Based on the real time identification of malicious flows, the ML model will be able to report insights to the Teraflow SDN controller at scale to perform security assessment.

A complementary use case consists of continuously assessing the data plane security status of the network across optical and IP layers. Initially, the monitoring data from optical and IP layers are performed independently by layer-specific attack detection and identification models. Then, security assessment is performed considering the monitoring data and security status of both layers. At optical level, the centralized cybersecurity component will leverage TeraFlow's equipment integration to use detailed physical layer monitoring parameters on the security assessment. For instance, Optical Performance Monitoring (OPM) data will be used to detect physical layer attacks on optical fibers.

At the control plane, a new use case addresses the design of ML algorithms that are deployed on top of the Teraflow SDN controller and react with resiliency to sophisticated adversarial AI attacks. This includes attacks to manipulate the information collected from the network, using the protocols in the control plane (e.g. telemetry or monitoring data). In particular, the set of ML components considered in Teraflow's ecosystem will be able to protect themselves against the recently appeared adversarial attacks that try to fool ML algorithms. Although adversarial attacks are still in their infancy, this use case will consider several well-known sources of attacks such as adversarial perturbations, out of distribution black-box attacks and white-box attacks.

B. Permissioned Distributed Ledger (PDL)

Blockchains have applications beyond cryptocurrencies and can be used for storing and processing data. For instance, data is not stored and processed in a central location; instead—a blockchain—which stores the data and the operations on the data, is copied and spread across multiple nodes, in which each node updates its blockchain to reflect a requested change, often by executing a smart contract. Consensus algorithms are used for agreeing on the blockchain state. TeraFlow will embrace the advantages of blockchains for network management. In particular, the TeraFlow OS will use blockchain technology to provide a trustworthy and resilient platform for storing, querying, and processing critical data about network resources and services owned and governed by different network entities. It will be privacy aware as well as transparent, resulting in an open, traceable, and fair sharing of network resources and services between stakeholders.

TeraFlow will deliver a permissioned distributed ledger that utilizes blockchains for network management. Furthermore, the network entities, their services, and the components of the TeraFlow OS will interact with the ledger through dedicated smart contracts. TeraFlow will provide a decentralized, robust, and trustworthy solution for storing, querying, and processing critical data for network resources and services. TeraFlow will contribute to blockchain technologies by providing research results on consensus algorithms and research on tools for analysing the security of smart contracts [10].

III. Conclusion and future work

Teraflow concept opens a unique opportunity to scale the SDN paradigm to Telco-grade. The downside is that manage securely requires envisioning potential security risks and prepare for them. Teraflow proposes a solution based on the extensive use of distributed resources in 5G (edge computing) and novel techniques in ML based on the telemetry capacity provided by SDN protocols. Furthermore, the Teraflow SDN controller will be equipped with resilient ML-based components in order to protect themselves against the so-called adversarial attacks.

Permissioned Distributed Ledgers (PDL) are expected to bring novel use cases to evolve security in B5G networks, such as smart-contracts, to enforce resource allocation or real time weaknesses analysis of network applications. The need to promote research on the security in introducing smart contracts is of extreme significance in order to provide personalized, multi-tenant B5G networks.

Acknowledgment

Work partially supported by the EC H2020 TeraFlow (101015857) and Spanish AURORAS (RTI2018-099178-I00).

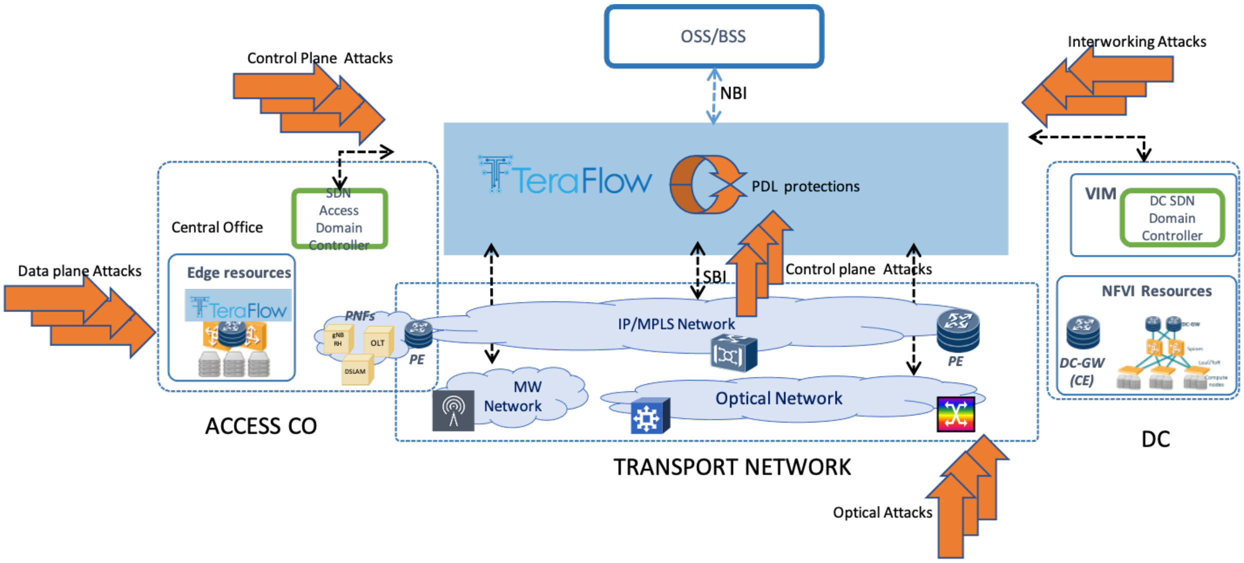

Fig. 1: TeraFlow architecture and attack vectors

References

[1] T. Dargahi *et al.*, "A survey on the security of stateful sdn data planes," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 3, pp. 1701–1725, 2017.

[2] "H2020 teraflow - secured autonomic traffic management for a tera of sdn flows," 2021. [Online]. Available: https://www.teraflow-h2020.eu/

[3] S. Lal *et al.*, "Nfv: Security threats and best practices," *IEEE Communications Magazine*, vol. 55, no. 8, pp. 211–217, 2017.

[4] C. Zhang, P. Patras, and H. Haddadi, "Deep learning in mobile and wireless networking: A survey," *IEEE Communications surveys & tutorials*, vol. 21, no. 3, pp. 2224–2287, 2019.

[5] M. Furdek *et al.*, "Optical network security management: requirements, architecture, and efficient machine learning models for detection of evolving threats [invited]," *IEEE J. Opt. Commun. Netw.*, vol. 13, no. 2, pp. A144–A155, 2021.

[6] T. Elsken, J. H. Metzen, F. Hutter *et al.*, "Neural architecture search: A survey." *J. Mach. Learn. Res.*, vol. 20, no. 55, pp. 1–21, 2019.

[7] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," *arXiv preprint arXiv:1406.2661*, 2014.

[8] Z. Lin, Y. Shi, and Z. Xue, "Idsgan: Generative adversarial networks for attack generation against intrusion detection," *arXiv preprint arXiv:1809.02077*, 2018.

[9] K. Eykholt, I. Evtimov, E. Fernandes, B. Li, A. Rahmati, C. Xiao, A. Prakash, T. Kohno, and D. Song, "Robust physical-world attacks on deep learning visual classification," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 1625–1634.

[10] M. Rodler *et al.*, "Sereum: Protecting existing smart contracts against re-entrancy attacks," *preprint arXiv:1812.05934*, 2018.